

451 CTAAATCAGAAAGTTACGAGCTGTTATGTATATGACATCCACAAATGC 500
 167 aaenglyilemetgluglyleugluproleuasnalaileargvalthr 184
 501 CAATGGTATTATGAAAGGCTTGGACACCTTGAATGCTATCAGAGTTAC 550
 184 lngluhalaaltpyrrtyrseraspasnalaiproilleserasnproasp 200
 551 AAGAGCGGTATGCTACTATTTGATATATGCTCTTAATTTCTAATCCAGAT 600
 201 gluserphelysarvglusergluserasnleuvalserthrserglnle 217
 601 GAAAGTTTTAAAGGAGTCGAAAGTAACCTGTTAGTACTTCTCAATT 650
 217 uSerleuMetArgGlnAlaLeuLysGlnLeuIleAspProAsnLeuAla 234
 651 ATCTTTGATCCGTCACACCTTTGAAGCACTGATTTGATCCCAATTTGGCA 700
 234 hrlsMetProLysGlnValProAspAspPheGlnLeuSerIlePheGlu 250
 701 CTAAATGCCAAACAAAGTCCGATGATTTTCAGCTAAGTATTTTGGAG 750
 251 SerGluAspLysGlnAspLysPylsTyrrasnLysGlyTyrglnasnleu 267
 751 TCTGAGGACAAAGGAGATTAATATTAATAAGATACCAAAATCTTTTGGAG 800
 267 rgllyGlyLeuValProThrLysProProThrProGlyAspProMetP 284
 801 TGTGTGTTTGTTCCTACTAACCACCACTCCAGAGACCCACCAATGC 850
 284 robProasnGlnProGlnThrThrSerValLeuIleargLysTyrrAlaIle 300
 851 CTCCAATCAACTCAAAACGACTTCAGTACTTATAGAAAGTATGCTATA 900
 301 GlyAspTyrSerLysLeuGluGlyAlaThrLeuGlnLeuThrGlyAs 317
 901 GGTGATTACTCTAAATTCCTTGAAGGTGCAACATTTACAGTTGACGGGGA 950
 317 PAsnValAsnSerPheGlnAlaArgValPheSerSerAsnAspIleGly 334
 951 TAACTGTAATAGTTTCAACGCGAGCTTTAGCAGTAATGATTTGGAG 1000
 334 LuArgIleGlnLeuSerAspGlyThrTyrrThrLeuThrGlnLeuAsnSer 350
 1001 AAGAAATTTGAATCATATCAGATGGAACCTTACTTTAAGTGAATCAATCT 1050
 351 ProAlaGlyTyrSerIleAlaGluProIleThrPheLysValGluAlaG 367
 1051 CCAGCTGGTTATAGTATCCGAGAGCCCAATCATTTTAAGGTTGAAAGCTGG 1100
 367 YLysValTyrrThrIleIleAspGlyLysGlnIleGluAsnProasnLysG 384
 1101 CAAGAGTATATCTATTTATGATGGAACCAAGATTGAAATCCCAATTAAG 1150
 384 LuIleValGluProTyrrSerValGluAlaTyrrasnAspPheGluLysPhe 400
 1151 AGATAGAGAGCCCTTACTACTAGTAAGCATATATGATTTTGAAGAAATTT 1200
 401 SerValLeuThrThrGlnAsnTyrrAlaLysPheTyrrTyrrAlaLysAnly 417
 1201 AGCGTTTAACTACACAAACATATGCAAAATTTTATATGCAAAATAATTA 1250
 417 sAsnGlySerSerGlnValValTyrrCysPheAsnAlaAspLeuLysSerP 434
 1251 AAATGGAAGTTCACAGGTGTCTATTTCTTTAATGCAGATCTAAATATCTC 1300
 434 robProAspSerGluAspGlyLysTyrrMetThrProAspPheThrThr 450
 1301 CACACAGCTCTGAAGATGTGGGAAACAAATGACTCCAGACTTACACA 1350
 451 GlyGluValLysTyrrThrHisIleAlaGlyArgAspLeuPheLysTyrr 467
 1351 GGAGAGTAAATAACACTCATATTTGACAGTCTGACCTTTTAAATATATAC 1400

467 rValLysProArgAspThrAspProAspThrPheLeuLysHisIleLysL 484
 1401 TGTGAACCAAGAGATACCGATCTCGACACTTTCTTAAACATATCAAAA 1450
 484 ysValIleGlnLysGlyTyrrArgGlnLysGlyGlnAlaIleGlyTyrSer 500
 1451 AAGTAATTTGAAAGGTTTACAGGAAAAAGGACACCTATTTGATATAG 1500
 501 GlyLeuThrGlnThrGlnLeuArgAlaAlaThrGlnLeuAlaIleTyrTy 517
 1501 GGTCTAAGTACAGACAAATTTGCTGGCGCTACTACAGTTACCAATATATTA 1550
 517 rPheThrAspSerAlaGluLeuAspLysAspLysLeuLysAspTyrHisG 534
 1551 TTTCACTGATAGTGTGAATATAGATAAGATAAATAAAGACTATACATG 1600
 534 LysPheGlyAspMetAsnAspSerThrLeuAlaValAlaLysIleLeuVal 550
 1601 GTTTTGGAGCATGAATGATATGATCTTTAGCAGTGTCTAAATCTTGTGA 1650
 551 GluTyrrAlaGlnAspSerAsnProGlnLeuThrAspLeuAspPheP 567
 1651 GAATACGCTTAAGATAGTAAATCTCCACAGCTAAGTACCTTATTTCTT 1700
 567 eIleProAsnAsnAsnLysTyrrGlnSerLeuIleGlyThrGlnTrpHisP 584
 1701 TATTCGAATTAACAATTAATATCAATCTTTATGGAACCTAGTGCATC 1750
 584 roGluAspLeuValAspIleIleArgMetGluAspLysLysGluValIle 600
 1751 CAGAGATTAGTTGATATTAATTCGTATGAGATATAAAGAAAGTATA 1800
 601 ProValThrHisAsnLeuThrLeuArgLysThrValThrGlyLeuAlaG 617
 1801 CCTGTAACTATATTTAATTTACATTGAGAAAAACCGTACTGCTTACTG 1850
 617 yAspArgThrLysAspPheHisPheGluIleGluLeuLysAsnAsnLysG 634
 1851 TGACAGAACTAAAGATTTCCATTTTGAATTTGAATTAATAATAATAC 1900
 634 lngluLeuLeuSerGlnThrValLysThrAspLysThrAsnLeuGluPhe 650
 1901 AAGAAATGCTTTCTCAAACTGTTAAACAGATTAACCAAACTCGAATTT 1950
 651 LysAspGlyLysAlaThrIleAsnLeuLysHisGlyGlnSerLeuThrIle 667
 1951 AAAGATGCTTAAAGCAACCATTTAATTTAAACATGGGAAAGTTTAACACT 2000
 667 uGlnGlyLeuProGlnLysTyrrSerTyrrLeuValLysGluThrAspSerG 684
 2001 TCAAGGTTTACCAAGAGTTATTTCTTACCTTGTCAAAGAAACAGATTCTG 2050
 684 lnglyTyrrLysValLysValAsnSerGlnLysValAlaAsnAlaThrVal 700
 2051 AAGGCTATTAAGTTAAAGTTAATAGCCAAAGATAGCAAAATGCTACAGT 2100
 701 SerLysThrGlyIleThrSerAspGluThrLeuAlaPheGluAsnAnly 717
 2101 TCAAAAAACAGAAATACAAATGATGACACACTTCTTTGAAAATATATA 2150
 717 sGluProValValProThrGlnLysAlaLysGlnLysIleAsnGlyTyrrLeu 734
 2151 AGAGCCTGTTGCTTCCTACAGAGTGTATCAAAAGATCAAAAGCTATACAG 2200
 734 laLeuIleValIleAlaGlyIleSerLeuGlyIleTyrPylIleHisThr 750
 2201 CTTTGATATGATATCGTGTATCAGTTTGGGATCTGGGGAATTCACAG 2250
 751 IleArgTleArgLysHisAsp 757
 2251 ATAGAGTAAAGAAACATGAC 2271

```

seq_name: /cgnl_7/ptodata/2/pna/US09_NEW_COMB.seq:us-09-494-297-3
seq_documentation_block:
; Sequence 3, Application US/09494297
; GENERAL INFORMATION:
; APPLICANT: PODBIELSKI, ANDREAS
; TITLE OF INVENTION: COLLAGEN-BINDING PROTEINS FROM STREPTOCOCCUS PYOGENES
; FILE REFERENCE: P06628USO/BAS
; CURRENT APPLICATION NUMBER: US/09/494,297
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2229
; TYPE: DNA
; ORGANISM: Streptococcus pyogenes
US-09-494-297-3

```

```

alignment_scores:
Quality: 1965.50      Length: 737
Ratio: 3.412          Gaps: 13
Percent Similarity: 78.155      Percent Identity: 54.410

```

```

alignment_block:
US-09-494-297-2 x US-09-494-297-3 ..

```

```

Align seq 1/1 to: US-09-494-297-3 from: 1 to: 2229

```

```

20 SerLysAsnSerLysArg.....PheThrValIhrIeuValIglValAlp 34
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
31 AGCGCTAACACAAACGACGACAAACGATCGGATTACTGAAAGTATT 80
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
34 eLeuMetIlePheAlaIeuValIhrSerMetValIglValAlp 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
81 TTGACGTTTCTACCTCGATAGAGATAGAGGTTTCTATCAGAGCGT 130
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51 heGlyLeuValGluSerSerThrProAsnAlaIleAsnProAspSer 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
131 TCGGAGCTGAAGAACAACTACCAAT.....AGCAACACTCAATT 174
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68 SerGluIyrArgTrpIyrGlyIyrGluSerIyrValArgGlyHisPro 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
175 CAAGATTATCCGTGATGCTATGATCTTATCTAAAGGCTACCCAGA 224
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
84 rTyLysGlnPheArgValAlaHisAspLeuArgValAsnLeuGlyHis 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
225 CTATAGCCCGTTAAAGACTTACATTAATTAAAGTAAATTTAGAGGAA 274
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
101 eArGSerIyrGlnValIyrCysPheAsnLeuLysValAlaPheProLeu 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
275 GTAAGGATTATCAAGCATACTGCTTATTAATTAACAAACATTTCCATCC 324
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
118 GlySerAspSerSerValLysIyrTrpIyrLysHisAspGlyIleSe 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
325 AAGTCACATAGCTTAAATGATCAATGATTAATAAACTTGAAAGAACTAA 374
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
134 rThLysPheGluAspIyrAlaMetSerProArgIleThrGlyAspGlu 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
375 TGAAGAACTTTATCAAGTTAGCAGATAAACCAAGATAGAGACGACAGCT 424
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
151 euAsnGlnLysLeuArgAlaValIleTyrAsnGlnHisProGlnAsnAla 167
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
425 TACAACAAATATATATGAGATCTCTATATATGATATCCATAATATCGGT 474
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
168 AsnGlyIleMetGluGlyLeuGluProLeuAsnAlaIleArgValIhrG 184
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
475 AATGGGATAATGAAGGATAGATCTCTAAACGCTATTTTAGTGACTCA 524
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
184 nGluAlaValIyrTrpIyrSerAspAsnAlaProIleSerAsnProAsp 201
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
525 AATGCTATTTGG...TATACTGATTCAGCTCAATTT...AATCGGATG 568

```

```

201 IuSerPheLysArgGluSerGluSerAsnLeuValSerThrSerGlnLeu 217
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
569 AAAGTTTAAACAGAGCTCGAAGTATGCTATTATGACACACAGCTTA 618
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
218 SerLeuMetArgGlnAlaLeuLysGlnLeuIleAspProAsnLeuAla 234
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
619 GCCTTAATGCAAAAGCTTAAAGAACTATGATCAAACTTAAGGGTTC 668
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
234 rLysMetCProLysGlnValProAspAspPheGlnLeuSerIlePheGlu 251
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
669 AAAATATTCCGAATTAACCTCCATCAGTTATCGGTTAACTGATTTGAT 718
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
251 eArgLysAspLysGlyAspLysTyrAsnLysGlyTyrGlnAsnLeuLeu 267
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
719 CTCATGAT.....AAGCCTTCCAAATCTTTTGAGT 750
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
268 GlyGlyLeuValProThrLysProProThrProGlyAspProPheMet 284
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
751 GCTGAGTATGTTCCGGAATACCTCCCAAAACGAGAGAGAG.....CC 794
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
284 oProAsnGlnProGlnIhrThrSerValLeuIleArgLysTyrAlaIleG 301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
795 TCCGCTAAAACTGAAAAAACATCATCATTAATCAGAAAAATATCGGAG 844
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 LysAspIyrSerLysLeuLeuGlnGlyAlaThrLeuGlnLeuThrGlyAsp 317
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
845 GTGAC...TCTAACTCTAGAGGAGCAACCTTAAGCTTCTCAATT 891
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
318 AsnValAsnSerPheGlnAlaArgValPheSerSerAsnAspIleGly 334
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
892 GAAGGAAGTGGTTTCAAGAAAAAGACTTTCAAGTAACTATTAGGAGA 941
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
334 uArgIleGluLeuSerAspGlyThrIyrThrLeuThrGluLeuAsnSer 351
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
942 AACTGTGAATTACCAATGAGGACTTATACCTTAACAGAAACATCATCTC 991
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
351 rOAlaGlyTyrSerIleAlaGluProIleThrPheLysValGluAlaGly 367
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
992 CAGATGATATTAATAATGCGGAGCGATTAAGTTAGTGTAAGATAA 1041
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
368 LysValIyrThrIle...IleAspGlyLysGlnIleGluAsnProAsn 383
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1042 AAAGTATTATTCGTCACAAAAAGATGTTCTCAAGTGGAAAAATCCAA 1091
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
383 sGluIleValGluProIyrSerValGluAlaIyrAsnAspPheGluIup 400
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1092 AGAAGTACGACAGCCATATCTCAGTGAAGCGTATATATGACTTTATG 1141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
400 heSerValLeuThr...ThrGlnAsnTyrAlaLysPheTyrTyrAlaLys 415
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1142 AAGAGTACTCTCGGGTTTACTCCATACGGAATAATCTATTAGCTTACA 1191
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
416 AsnLysAsnGlnLysSerSerGlnValValIyrCysPheAsnAlaAsp 432
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1192 AATAAGGATTAAGTTCACAAAGTTGTCTACTGCTTCAATGCTGATTACA 1241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
432 sSerProProAspSerGluAspGlyGlyLysThrMetThrProAspPhe 449
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1242 CTCACACCTGACATCATATGATAGGTGAGACTATAATACAGACTACTA 1291
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
449 hrThr...GlyGluValLysTyrThrHisIleAlaGlyArgAspLeuP 464
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1292 GTACGATCAAAAGATCAAGTACACACATACGCGAGGTAGACTTGT 1341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
465 LysTyrThrValLysProArgAspThrAspProAspThrPheLeuLys 481
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1342 AAATATGCGCTTAAGACCGAGAGATACAAATCCAGAAAGCTTTAAAGCA 1391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
481 sIleLysLysValIleGluLysGlyTyrArgGluLysGlnAlaIleG 498
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1392 CATTAATAAAAGTAATTAAGAAAGCTCAAGAAAAAGGATATGAC.... 1437
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
498 IuTyrSerGlyLeuThrGluThrGlnLeuArgAlaIleThrGlnLeuAla 514

```


4078 GTCACATAAAGACAAATACGAACTGCAACT.....GCAGATCAG 4038
 194 AlaProIleSerAsnProAspGluSerPheIysArgIleuSerGluSerAs 210
 4037 CCTAAAGTGCATAGTACTAGTACCAAGTGTAAAGAACTAGTACTAA 3988
 210 nIeuValSerThrSerGlnIleuSerIleuMetArgIleAlaIleuIysGln 227
 3987 CATGCATTCACACCAAAACGCTACAGCTAATCAATCT..... 3951
 227 euIleAspProAsnIleuAlaThrIysMetProIysGlnValProAspAsp 243
 3950ACTACAAAACCTGCATGTACACAAACAAGAT 3918
 244 PheGlnIleuSerIlePheGluSerGluAspIysGluSplysIyTrAsnIy 260
 3917 AAATCATCACTACATAGTATAGTAAGAACT.....GATAAA..... 3882
 260 sGlyTyArgIleAsnIleuSerGlyIyGluValProThrIysProIle 277
 3881AGTAATTTAACACAGCAAAAGATGTTTCACTACACCTTAAA 3839
 277 hrProGlyAspProIleProMetProProAsnGlnProGlnThrThrSerVal 293
 3838 CAACGACTATTAAACCAAGAACTTTAAATCGCATGCGACGTAATACGT 3789
 294 IeuIleArgIySTyAlaIleGlyAspTySerIysIleuIleuGluIyAl 310
 3788 GCAGCTCCACA..... 3777
 310 aThrIleuGlnIleuThrGlyAspAsnValAsnSerPheGlnAlaArgVal. 326
 3776CAAGAACAAATGTTAAATGAT.....AAAGTAC 3749
 327 .PheSerSerAsnAspIleGlyIleuArgIleGluIleuSerAspGlyThr 342
 3748 ATTTTCAAATATATGACATTCGCGATGATAAGCAATGTTAAATCAGACT 3639
 343 TyThrIleuThrGlu.....IleuAsnSerPr 351
 3698 ACTGCTAAACTGAAATTTTGGCAACTTCAAGTGATGTTTAAATTTAAA 3649
 351 oAlaGlyTySerIleAlaIleuProIle.....ThrP 362
 3648 AGCAAAATTTACCAATCGATGATCTGTAAAGAGCGGACATACATTCTT 3599
 362 heIys.....ValGluAlaGlyIysValTyThrIleIleAsp 374
 3598 TTTAAATTTGGTCAATATTTCCGTCAGAGATCAGTAAGATTAACCTTCA 3549
 375 GlyIysGlnIleGluAsnProAsnIysGluIleValGluProTySerVa 391
 3548 ACTCAAAATTTATATATGCCCAAGTAATATTTAT.....GCCAA 3508
 391 IgluAlaTyAsnAspPheGluIleuPheSerValIleuThrThrGlnAsn 408
 3507 AGGTATTATGATAGTACAAACAACACACACATATATCTTTAGCAACT 3458
 408 yAlaIysPheTyTyr.....Tyr 413
 3457 ATGTAGATCATATATACAAATGTTAGAGTACCTTTGAACAAGTTGCATTT 3408
 414 AlaIysAsnIysAsnGlySerSer..GlnValValTyrcysPheAsnAl 429
 3407 GCGAAAGCTTAAATGCACAACATGATAAAACAGCTTATAAATGGAAGT 3358
 429 aaSPleuIysSerProProAspSerGluAsp.....Glyc 441
 3357 AACTTATGATATGATATACATATAGCAAGAAATCATGTCGATATATGTA 3308
 441 IyIysThrMetThrProAspPheThrThrGlyGluValIySTyThrHis 457
 3307 ATTAATAAAGCAACACCGCTTATTTCAAGTACAAACATATATTAACAATGA 3258

458 IleAlaGlyArgAspIleuPheIySTyThrValIysProArgAspThr.. 473
 3257 GATTATTCGCTAATATGACTGCATATGTAATACACCTTAAAAATACATA 3208
 474 .AspProAspThrPheIleuIysHisIleIysIysValIleGluIysGlyT 490
 3207 TACTAAACAAACGTTGTACTAATTTA.....ACTGAT 3173
 490 yTrArgIuIysGlyGlnAlaIleGluTySerSerGlyIleuThrGluThrGln 506
 3172 ATAAATTTAATCCAAATGCAAAAAACCTCAAA..... 3141
 507 IeuArgAlaAlaThrGlnIleuAlaIleTyTyThrThrAspSerAlaGlu 523
 3140ATTACGACAGTACAGATCAAAATCA 3115
 523 uIleAspIysAspIysIleuIysAspTyHisGlyPheGlyAspMetAsn 540
 3114 ATTT..... 3111
 540 sPserThrIleuAlaValAlaIysIleIleuValGluTyAlaGlnAspSer 556
 3111 3111
 557 AsnProGlnIleuThrAspIleuAspPheIleProAsnAsnAsnIy 573
 3110GTGATAGTTTCACCCCTGATCTCAAA 3082
 573 sTyArgIleuIleGlyThrGlnTyTrHisProGluIysPheValAspI 590
 3081 ACTTAAAGATGTT.....ACTGATCAATTCGATG 3053
 590 IeIleArgMetGluAspIysIysGluValIleProValThrHisAsnIeu 606
 3052 TTATTTATAGTAATGATAATPAAA.....ACAGCTACAGTC 3018
 607 ThrIleuArgIySTyValThrGlyIleuAlaGlyAspArgThrIysAsp 623
 3017 GATTTAATGAAAGGCCAAACA..... 2997
 623 eHisPheGluIleGluIleuIysAsnAsnIysGlnIleuIleuSerGln. 639
 2996ACGACAAATTAACATACATCATCTTCAACAG 2966
 640Thr 640
 2965 TTGCTATTCAGATATAGTTCACACAGATAATGAAAAATGATTATACT 2916
 641 ValIysThrAspIysThrAsnIleuGluPheIysAspGlyIysAlaThrI 657
 2915 TTAGACACTGCACAAACCTAAATATAGTTGGTCAAAATGATTATCCAAATGT 2866
 657 eaSPleuIysHisGlyIleuSerIleuThrIleuGlnGlyIleuProGluIyT 674
 2865 GAAT.....GCCTATCATACTGCTAATGGGACCAAAAGAAAT 2828
 674 yTrSer.....TyIleuValIysGluThrAspSerGluIyTyIys 687
 2827 ATATATCAGAGTACTATGATGGAACATACAAATTAAGATGCT..AAA 2781
 688 ValIysValAsnSerGlnIleuAlaIleAsnAlaThrVal.....Se 701
 2780 CAAGATCCCATGAAAGGATTAAGGCTGTTATGTCATTTCTTAAGA 2731
 701 rIySTyThrGlyTyIleThrSerAspGluThrIleuAlaPheGluAsnAsnIys 717
 2730 TAGTACGCTTAAGAAATTAATGATCTACACACAGATGAAAATGTAATA 2682
 seq_name: /cgnl_7/ptodata/2/pna/US08_NEW_COMB.seq:US-08-956-171C-472
 seq_documentation_block:
 ; Sequence 472, Application US/08956171C

GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gill H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171C
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Hoover, Kenley K.
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 610-5790
FAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 472:
SEQUENCE CHARACTERISTICS:
LENGTH: 6806 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 472:
US-08-956-171C-472

alignment_scores:
Quality: 136.50 Length: 794
Ratio: 0.403 Gaps: 39
Percent Similarity: 42.695 Percent Identity: 19.270

alignment_block:
US-09-494-297-2 x US-08-956-171C-472/rev ..

Align seq 1/1 to reverse of: US-08-956-171C-472 from: 1 to: 6806

```
137 PhgGluAspIyrAlaMetSerProArgIleThrGlyAspGluLeuAsnG1 153
|||||
6573 TTTAATGATTATGACGCT.....CGTATT.....ATGAACTTAAAC 6536
|||||
153 nLysLeuArgAlaValMetYrAsnGlyHisProGlnAsnAlaAsnGlyI 170
|||
6535 AAAAGATAAAACTTA.....GATGGTATGATA 6507
|||||
170 LemetGluGlyLeuGluProLeuAsnAlaIleArgValThrGlnGluAla 186
|||
6506 TTGCTGAAGTTTAAACAGCTGTGCTGCTGCTATTCAGAAAGAAATTA 6457
|||||
187 ValTrrpYrIyrSerAspAsnAlaProIleSerAsnProAspGluSerPh 203
|||
6456 TTG.....CAATT 6449
```

```
203 eLysArgGluSerGluSerAsnLeuValSerThrSerGlnLeuSerLeuM 220
|||||
6448 TGAAGGACCAACGAATCTAAATTG...GGTACTTCTGAA..... 6412
|||||
220 eLArgGlnAlaLeuYrGlnLeuIleAspProAsnLeuAlaThrLysMet 236
|||||
6411.....GCTAGAAAGCTGTGTGATTCACTTGTGTGACACAAATTG 6373
|||||
237 ProLysGlnValProAspAspPheGlnLeuSer..... 247
|||||
6372 CCATTCTATTTAGAAGAAAAGCAATGCTTAATCACTTGTGA AAAA 6323
|||||
248 .....IlePheGluSerG 252
|||||
6322 AGCGATTAAAGCACAAACGAAGGAGAGCTGCACGTAAAGCTCGAAG 6273
|||||
252 LysAspLysGlyAspLysTyrAsnLysGlyTyrGlnAsnLeuSerGly 268
|||||
6272 ATGCTGCTTCAGGTAAGAAAACAGCGTAAACACACTTGTCTATCTGCT 6223
|||||
269 GlyLeuValProThrIyrSerProProThrProGlyAspProPhePro 285
|||||
6222 AATTAAACCTGCACAA.....AGTAA 6200
|||||
285 oAsnGlnProGlnThrThrSerValLeuIleArgLysTyrAlaIleGly 302
|||||
6199 AAACACTGAAAATAATGAAATGATTGATTAGTCGAAGTGATTCGCGAG 6150
|||||
302 sPrrYserLysLeuLeuGly.....AlaThrLeuGlnLeu 314
|||||
6149 GTTCAGCAAACTTGACGACGACCGCAATTCACAGCATATTACCATTA 6100
|||||
315 ThrGlyAspAsnValAsnSerPheGlnAlaArg.....ValPheSe 328
|||||
6099 CGTGTGAAGTATTATATACAGAAAGCAGCTGTAAGATATTTTTAA 6050
|||||
328 rSerAsnAspIle.....Glyc 334
|||||
6049 AAATGAAGAAATTAATATATATCCACACAAATCGGGCAGCGCTTGGA 6000
|||||
334 LysrGlyLeuLeuSerAspGlyThrThr.....LeuThr 346
|||||
5999 CTGACTTTAAATGAAAGATGATTAATATATATCGTGAATTATTAAGCT 5950
|||||
347 GluLeuAsnSerProAlaGlyTyr..... 354
|||||
5949 GATGCTGATCTGATGCTGCGCATATTCAGAGTGTATTTACATTCTT 5900
|||||
355 .....SerIleAlaG 358
|||||
5899 CTTCAAATATATGAAACCGCTGTCTTCAGCAGGTCGTATTATTGCTT 5850
|||||
358 LysProIleThrPheLysValGluAlaGlyLys..... 368
|||||
5849 TACCTCCACTTTATTAATTTGAAAAGTAAAGGCAAAACAAACCGAGTT 5800
|||||
369 ValTyrThrIleLeuAspGlyLysGlnIleGluAsnProAsnLysGluI 385
|||||
5799 GAATACCGCTTGACAGAC...GAAAGCTTAAATAATTCAGAAAAGAACT 5753
|||||
385 eValGluProTyrSerValGluAlaTyrAsnAspPheGlnGluPheSerV 402
|||||
5752 TGGTAAAGGCTTCACGTTACAGCTTAAGAGGTTTGGTGAA..... 5710
|||||
402 alleu..ThrThrGlnAsnTyrAlaLys..... 410
|||||
5709 ..ATGAACCTTGAAACAAATTATGGAAACGACGATGAACCAAGAAACGA 5662
|||||
411 .....PheTyrAlaLysAsnLysAsnGlySerSerGlnVal. 423
|||||
5661 ACTTATATGCTGTACAAAGTTGAAGATGAAGTGCGTCACTTAACGCTG 5612
|||||
424 .....ValTyrCysPheAsnAlaAspLeuLysSerProPro 435
```


[illegible]


```

1448 ATATCAACATTTTCGATACATTAATGACATTCATACAGAAAATGTA 1497
      ::::::::::: :::: ::::::::::: |||
613 ThrclyleuAlaGlyAspArgThrlYsAspHehIsPheGluIleGluLe 629
      ::::::::::: ::::::::::: :::: |||:
1498 GCATTTTCTTCT.....GAATTAACACTGATCAAGAAAT 1532
      ::::::::::: ::::::::::: :::: |||:
629 uLYsAsnSnLysGln.....GluLeuLeuSerGlnThrValLysThr 643
      ::::::::::: ::::::::::: :::: |||:
1533 AAAAGAATTCAAAATGTCACCTCAAAATCTATTCTGACACTAGAAAATG 1582
      ::::::::::: ::::::::::: :::: |||:
644 .....AspLysThrAsnLeuGluPheLysAspGly... 653
      ::::::::::: ::::::::::: :::: |||:
1583 CAATACTAAGTCATTTGATTTAACTAAATTA.....AAAGACGTCCT 1626
      ::::::::::: ::::::::::: :::: |||:
654 .....LysAlaThrIleAsnLeuLysHisGlyLysSerLeuThrLeuGl 668
      ::::::::::: ::::::::::: :::: |||:
1627 ACATTTAAATTTTAATGATCTCAACACAGAACAAATATTAACTTTACT 1676
      ::::::::::: ::::::::::: :::: |||:
668 nGlyLeuProGluGlyTyrSerTyrLeuValLysGlu.....Thra 682
      ::::::::::: ::::::::::: :::: |||:
1677 AGGCTTAAGTAAACCCCTTCAGTCTCTAAACCTGAAAACCAAAAGATC 1726
      ::::::::::: ::::::::::: :::: |||:
682 sPserGluGlyTyrLysValLysValAsnSerGlnGluValAlaAsnAla 698
      ::::::::::: ::::::::::: :::: |||:
1727 AAAGTTGGAACCCACAAACAGATACCTCTACTCAAAAACAGAAAGCGCA 1776
      ::::::::::: ::::::::::: :::: |||:
699 ThrValSerLysThrGlyIleThrSerAspGluThrLeuAlaPheGluAs 715
      ::::::::::: ::::::::::: :::: |||:
1777 ACAAGTTCA.....ACAGATTCACAAACAAAGCTCAACACTGA 1811
      ::::::::::: ::::::::::: :::: |||:
715 nAsnLysGluProValAlaProThr 723
      ::::::::::: ::::::::::: :::: |||:
1812 AAACCAAAACCGCGCTGACAAACA 1836
      ::::::::::: ::::::::::: :::: |||:

seq_name: /c9n1_7/ptodata/2/pna/us09_NEW_COMB.seq:us-09-430-590E-139

seq_documentation_block:
; Sequence 139, Application US/09430590E
; GENERAL INFORMATION:
; APPLICANT: Poultier, et al.
; TITLE OF INVENTION: UNUSUAL RETROTRANSPON FROM THE YEAST CANDIDA ALBICANS
; FILE REFERENCE: 674521-2001.1
; CURRENT APPLICATION NUMBER: US/09/430,590E
; CURRENT FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/106,342
; PRIOR FILING DATE: 1998-10-30
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 139
; LENGTH: 9850
; TYPE: DNA
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(9850)
; OTHER INFORMATION: 'n' can be any nucleotide 'a', 'c', 'g' or 't'
US-09-430-590E-139

alignment_scores:
  Quality: 131.50      Length: 635
  Ratio: 0.476        Gaps: 33
  Percent Similarity: 43.465  Percent Identity: 18.583

alignment_block:
US-09-494-297-2 x US-09-430-590E-139
Align seg 1/1 to: US-09-430-590E-139 from: 1 to: 9850

75 TyrGluSerTyrValArgGlyHisProTyrTyrLysGlnPheArgValAl 91
||| ::::::::::: |||
3175 TATGCGTTTCATATCCGAAATCATACACTATATAAAGAAATTT..... 3216

```

```

91 aHisAspLeuArgValAsnLeuGluGlySerArgSerTyrGlnValTyrC 108
      ::::::::::: ::::::::::: :::: |||
3217 .....GATGGTCTACTCTTATAGACTTAC. 3243
      ::::::::::: ::::::::::: :::: |||
108 ySPheAsnLeuLysLys.....AlaPheProLeuGlySerAsp 120
      ::::::::::: ::::::::::: :::: |||
3244 ..TATGCTTTATCTTAATATAGCTCATACCATTTTTCATTTGGAAACGAC 3291
      ::::::::::: ::::::::::: :::: |||
121 SerSerValLysLysTyrTyrLysLysHisAspGlyIleSerThrLysP 137
      ::::::::::: ::::::::::: :::: |||
3292 GTTTTGATA...AAATGTCYAGTGTACAGAACGCTATTTCATTAACACT 3338
      ::::::::::: ::::::::::: :::: |||
137 e.....GluSprYrAlaMetSerProArgLethrGly..... 148
      ::::::::::: ::::::::::: :::: |||
3339 ACCATCTTCAAGAGATTAAGCTTTTCTTACAGTGTATGTTGCTCTTTC 3388
      ::::::::::: ::::::::::: :::: |||
149 .....AspGluLeuAsnGlnLysLeuArgAlaValMetTyr 160
      ::::::::::: ::::::::::: :::: |||
3389 TCGGTTAGCGCTCAGATTCCTTACCTTCAGATT..TTAGTTTCCACG 3435
      ::::::::::: ::::::::::: :::: |||
161 AsnGlyHisProGlnAsnAlaAsnGlyIleMetGluGlyLeuGluProLe 177
      ::::::::::: ::::::::::: :::: |||
3436 AAAGATATCCAGTTATTACAA.....TCAACATCCGCCAAT 3476
      ::::::::::: ::::::::::: :::: |||
177 uAsnAlaIleArgValIThrGlnGluAlaValIThrTyrTyrSerAspAsn 194
      ::::::::::: ::::::::::: :::: |||
3477 AGCGCAGATGCAAGTCAATCAATGACTATTGTGCATATCAATATCGGAAT 3526
      ::::::::::: ::::::::::: :::: |||
194 lAProIleSerAsnProAspGluSerPheLysArgGluSerGluSerAsn 210
      ::::::::::: ::::::::::: :::: |||
3527 GCTCATATAAGCTAT..GACGATACATTC..... 3552
      ::::::::::: ::::::::::: :::: |||
211 LeuValSerThrSerGlnLeuSerLeuMetArgGlnAlaLeuLysGluLe 227
      ::::::::::: ::::::::::: :::: |||
3553 .....TATACACTTT 3563
      ::::::::::: ::::::::::: :::: |||
227 uLLeAspProAsnLeuAlaThrLysMetProLysGlnValProAspAsp 244
      ::::::::::: ::::::::::: :::: |||
3564 GAATCACCACCAATGATTCGCACAAAC..... 3588
      ::::::::::: ::::::::::: :::: |||
244 heGlnLeuSerIlePheGluSerGluAspLysGlyAspLysTyrAsnLys 260
      ::::::::::: ::::::::::: :::: |||
3589 .....CAACATGATACCGTGCAGCAATATAATATCTC 3621
      ::::::::::: ::::::::::: :::: |||
261 GlyTyrGlnAsnLeuLeuSerGlyGlyLeuValProThrLysProProThr 277
      ::::::::::: ::::::::::: :::: |||
3622 GAATATGAAAC.....CGTCCAAATGT 3644
      ::::::::::: ::::::::::: :::: |||
277 rProGlyAsp...ProPrometProProAsnGlnProGlnThrThrSerV 293
      ::::::::::: ::::::::::: :::: |||
3645 ACCATTTGAATATCATGTGTAACCTCTCGTACAAATTCATGACGAGGAA 3694
      ::::::::::: ::::::::::: :::: |||
293 alLeuIleArgLysTyrAlaIleGlyAspTyrSerLysLeuLeuGluGly 309
      ::::::::::: ::::::::::: :::: |||
3695 TTATCGATCGACCA.....GATATTAGACCTAGAGCTGATCC 3732
      ::::::::::: ::::::::::: :::: |||
310 AlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheGlnAlaArgVa 326
      ::::::::::: ::::::::::: :::: |||
3733 ACCTGGCAACGTATGCCGTGATGCCAACATATCATCAG...GAACAAACAC 3779
      ::::::::::: ::::::::::: :::: |||
326 lPheSerSerAsnAspIleGlyGluArgIleGluLeuSerAspGlyThrT 343
      ::::::::::: ::::::::::: :::: |||
3780 TGTACAGACTCTGTATCAATGAGGAGTTAGATACCATGATCAACACGAA 3829
      ::::::::::: ::::::::::: :::: |||
343 yTrhLeu..... 345
      ::::::::::: ::::::::::: :::: |||
3830 ACCAATCAACAGATCTGGGAGGATATTACCCGGGCAACAGGTGCC 3879
      ::::::::::: ::::::::::: :::: |||
346 .....ThrGluLeuAsnSe 350
      ::::::::::: ::::::::::: :::: |||
3880 ACCGATATTATTGGCAATTTCAGATGCGGCGGCTACCACTCTAAACAC 3929
      ::::::::::: ::::::::::: :::: |||

```



```

; LENGTH: 666448
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-12214

```

```

alignment_scores:
    Quality: 130.50      Length: 720
    Ratio: 0.384         Gaps: 32
    Percent Similarity: 47.222    Percent Identity: 18.333

```

```

alignment_block:
US-09-494-297-2 x US-09-335-032-12214

```

```

Align seg 1/1 to: US-09-335-032-12214 from: 1 to: 666448

```

```

52 G1yLeuValG1uSerThrProAsnAla1leAsnProAspSerSe 68
    ||| :|||:||||| :|||:|||||:|||||:|||||:|||||:
296795 GGTTCATTAAGGCGCTCAATGTTCTATATGACGCGGAGCACTAC 296844
68 r.....GluTyrAgtTPTyrglyTyrglySerTyra 80
    ||| :|||:||||| :|||:|||||:|||||:|||||:
296845 TCAGCCCGTTACAAATTAAGATTAAGGATTAATCAATCCATCTAG 296894
80 rgsly..... 81
296895 GTGGTGAAGATTACAAATGATATGGCAGTTGATAGAAATAGTTT 296944
82 .....HisProTyrTyrglyGlnPheArgValAlaHisAspleuAr 95
    ||| :|||:||||| :|||:|||||:|||||:|||||:
296945 TTGGAAACACACCCAGCCATAGCA.....ACTGATGAAATGCA 296982
95 gValAsnLeuGluGlySerArSerTyrglyValTyrglyCysPheAsnLeu 112
    ||| :|||:||||| :|||:|||||:|||||:|||||:
296983 CGCTAATCCCAAGGCTTACCAAAATCAACCAAGCAGACAGAGCA 297032
112 yslAsnAlaPheProLeuGlySerAspSerValTyrglyStrPTyrgly 128
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297033 AGTTAATTTTAAGGCGCAATTCGAGCAAGTAT..... 297067
129 LysHisAspGlyLeuSerThrLysPheGluAspTyrglyAlaMetSerProAr 145
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297068 .....AACATAGAAATCATCATCAACGATATGATTTCCCTACTTC 297108
145 gileThrGlyAspGluLeuAsnGlnLysLeuArgAlaValMetTyrAsnG 162
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297109 TATTAACAGACAGGAATTCAGAAATTTAT..... 297139
162 LysHisProGlnAsnAlaAsnGlyLeuMetGluGlyLeuGluProLeuAsn 178
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297140 .....GCAGACTCGTTATTTGACATTTGCAAAACCATTAAT 297175
179 AlalLeuArgValThrGlnGlnAlaValTPTyrglySerAspAsnAlaPr 195
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297176 GAC...GCTGTACAAAACAA...TTGCGTGGCTATGTGAACAATTTACC 297219
195 oileSerAsnProAspGluSerPheLysArgGluSerLysSerAsnLeuY 212
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297220 T.....GACATTAATGGCG 297233
212 aL.....SerThrSerGlnLeuSerLeuMetArgGlnAlaLeu 224
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297234 TCATTTTGGCGGAGGCTCTCCGGAATCCCATTTGCAAGATCAATTA 297283
225 LysGlnLeuIleAspProAsnLeuAlaThrLysMetProLysGlnValPr 241
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297284 ATCAAACTCTATCCGAAGAA.....AAAGCTTGAGAAATGTCAA 297324
241 oAspAspPheGlnLeuSerThrPheGluSerGluAspLysGlyAspLysT 258
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297325 TCGTGAT...GAATCAAGCTGTGAATGGTGTGTATGAGAGGATCAAGT 297371
258 yAsnLysGlyTyrglyAsnLeuLeuSerGlyLeuValProThrLys 274

```

```

297372 TATCTAATCGTTAAG.....ACCAAG 297394
    ||| :|||:||||| :|||:|||||:|||||:|||||:
275 ProProThrProGlyAspProPromeLePro.....Pr 285
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297395 CCGTTAAATGTTGTTGACCGTCTGTAAATACCTTATTCATAAATTATC 297444
285 oAsnGlnProGlnThrSerValLeuIleArgLysTyraAla1leGlyA 302
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297445 AAACGAATCTGAACGTATATGTCTTACCGCGGAGAGCTTATCCAA 297494
302 sPTyrSerLysLeuLeuGluGlyAlaThrLeuGlnLeuThrGlyAspAsn 318
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297495 ACAAAACATCTATTTGACAAACAGCACT.....GATTCG 297529
319 Val...AsnSerPheGlnAlaArgValPheSerSerAsnAsp1leGly 334
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297530 ATTCCTAATTAATTTTACCAATGATTAATTTGAGAAATGTTAAATGTTGCA 297579
334 uArgileGluLeuSerAspGlyThrTyrrThrLeuThrGluLeuAsnSerP 351
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297580 AACTATACAGTTAATTCAGAGACTATAAAGAAATTCATATTCCTGATA 297629
351 roAlaGlyTyrSerIleAlaGluProIlePhePheLysValGluAlaGly 367
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297630 AGTGTCTGACAGAGTTCGCTAATACATTAATTCGCTTCAGATGAT 297679
368 LysValTyrrThrIleLeasp..... 374
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297680 AGATTAATCTCTATTCAGAGGTTAACTGCAATTCGACAGCAAAATGA 297729
375 .....GlyLysGlnIleGluAsnProAsnLys..... 383
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297730 CATACGTAATCTCAGCAATTTAAGAAACAAAGCAGCGCTTGGCTTTA 297779
384 .....GluIleValGluProTyrSerValGluAlaTyraAsn 395
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297780 CTTCGAGATGTTGATGATCAAAAGGCTTCTCCTCAGAACCTTGCGGT 297829
396 AspPheGluLeuPheSerValLeuThrThrGlnAsnTyrrAlaLysPheTy 412
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297830 TTGCATAGATATCAAGTGTCTCGATTAACAGATTAAGCAAGATTTCA 297879
412 rTyrrAlaLysAsnLysAsnGlySerSerGlnValTyrglyCysPheAsnA 429
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297880 ATTCAGAAAGAAATTTAAACCTTCTGAAATTAATCTGTATAGATGATGAA 297929
429 laAspLeuLysSerProProAspSerGluAspGlyLysThrMetThr 445
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297930 ACCTGCTAATGATGATGAATGATGAATGACCAAAATGCCAAGTA 297979
446 ProAspPheThrThrGlyValLysTyrrThrHisIleAlaLysArgAs 462
    ||| :|||:||||| :|||:|||||:|||||:|||||:
297980 GAAGAGTTATCG.....GAGATGCTTAAGTAAATTTGATTTGCTCGCA 298023
462 PleuPheLysTyrrThrValLysProArGAspThrAspProAspThrPheL 479
    ||| :|||:||||| :|||:|||||:|||||:|||||:
298024 AGATGCATCCTT.....GATCGACCCGAGAGATATAG 298058
479 euLysHisIle.....LysLysValIleGlu..... 487
    ||| :|||:||||| :|||:|||||:|||||:|||||:
298059 TTAGCAGAAATTAAGAAATTTGAATATTAATAAAGAAATAGCACTTAC 298108
488 .....LysGlyTyrrAr 491
    ||| :|||:||||| :|||:|||||:|||||:|||||:
298109 ATGGAATTCGCAAGAACCTTTGAACCTCAACAATTTAAAGAAATGCT 298158
491 gGluLysGly.....GlnAlaIleGluTyrrSerGlyLeuThrG 504
    ||| :|||:||||| :|||:|||||:|||||:|||||:
298159 TGAAGAAAGCCATTAAGTTACTTCAAGCTATAGAAACCCATTAAGAAATACCG 298208
504 LuThrGlnLeuArgAlaAlaThrGlnLeuAlaIleTyrrTyrrPheThrAsp 520
    ||| :|||:||||| :|||:|||||:|||||:|||||:

```

```
298209 TTGAAGATTTTGGATCAATTTGAAACCGAGTTGGCGATACATAGAT 298258
521 SerialagluLeuAspLysAspLysLeuLysAspLysGlyPheGlyAs 537
298259 AATGTTAGAGAGAAATTTTAAAGATTAAG..... 298288
537 pmetasnaspSerThrLeuAlaValAlaLysIleLeuValGluTyrAlaG 554
298289 .....CACACGCGTATGTCGACAGCGCTTATCTACATGGGAGG 298328
554 In.....AspSerasnProGluLeuThrAspLeuAspPhePhele 568
298329 AACCTTAACCTCTTTTAAATTCATTAAGGAAATAGAGAGTTCTGT 298378
569 ProAsnAsnAsnLysTyrGlnSerLeuIleGlyThrGlnTTPHSPGCI 585
298379 GCAGAAAGC.....CTAATGGCGAAGACCTTCGT...GA 298410
585 uAspLeuValAspLeu.....IleArgMetG 594
298411 ACATTTATTTGAAATCAATTAATTAATTTGATATGTCGAGAACTAG 298460
594 LuAspLysLysGluValIleProValThrHisAsnLeuThrLeuArgLys 610
298461 AGGAAAGAACTGCGTTATATAA..... 298483
611 ThrValThrGlyLeuAlaGlyAspArgThr.....LysAs 622
298484 .....ACGCGTGATGAAAGTCGCTTAATGAATGAAGAA 298518
622 PpheHisPhe...GluIleGluLeuLysAsnLysGlnGluLeuLeus 638
298519 GTTACATTTAAGAACTTCGCCCTACAAAGAGAGAGGAAAGCTGA 298568
638 erGlnThrVal.....LysThrAspLys 645
298569 AAGAAAGAACTTGACAGAGAAAGCAGAAACAAATGAAGAAACAAATCG 298618
646 ThrAsnLeuGluPheLysAspGlyLysAlaThrIle...AsnLeuLysHi 661
298619 ACAGTAATCAACCTGCGCTGACGATTAAGTAATCTATTTGCAATGAC 298668
661 sGlyGluSer 664
298669 CACCGAGTCG 298678

seq_name: /cgnl_7/prodata/2/pna/us08_NEW_COMB.seq:us-08-956-171C-268

seq_documentation_block:
; Sequence 268, Application US/08956171C
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; G11 H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956, 171C
```

```

; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoover, Kenley K.
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 610-5790
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 268:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4702 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 268:
US-08-956-171C-268

alignment_scores:
Quality: 129.50 Length: 722
Ratio: 0.363 Gaps: 31
Percent Similarity: 49.446 Percent Identity: 18.837

alignment_block:
US-09-494-297-2 x US-08-956-171C-268/rev ..
Align seg 1/1 to reverse of: US-08-956-171C-268 from: 1 to: 4702

14 AsnThrGlnArgValLeuSerLysAsnSerLysArgPheThrValThrIle 30
2170 AACACAGAGAAAGTTGTAAGTAATGTACTTAACGACTTAAACAAAACTT 2121
30 uValGlyValPheLeuMetIlePheAlaLeuValThrSerMetValGlyA 47
2120 AGTCGAAATTTATTTAACA.....ACGATCATGTGATTAATGATTAATG 2077
47 lAluThrValPheGlyLeu.....ValGluSer 56
2076 CACAAAAAATGTTGGCGCTATATGACGCGTGAACATGCTGTGATAGT 2027
57 SerThrProAsn...AlaIleAsnProAspSerSerSerGluTyrArgTr 72
2026 AATTCCTCGAATTTACTTTAATCA..... 2000
72 pTyrGlyTyrGluSerTyrValArgGlyHisProTyrTyrLysGlnPheA 89
1999 .....ATTACGACTCCCGGAATTAATTACAGATA 1969
89 rGValAlaHisAspLeuArgValAsnLeuGluLysSerArgSerTyrGln 105
1968 CGCTTGTAATTCATTTCTGCACAAACAAACATTACAAATGCTCCAA 1919
106 ValTyrCysPheAsnLeuLysAla.....PheProLeuG1 118
1918 ACATACATAATATCTTATTGAGTCGGAATTCAGATACGTTCAAGGTAA 1869
118 ySerAspSerSerValLysLysTyrTyrLysHisAspGlyIle... 133
1868 CACAGATATATAGTTGCTGACTTAATGTGAACAAACAAATCATATATTG 1819
133 ..... 133
1818 ACGAGCACAATACAGCGATGATAAATGTTACAGATTAATAAATCGCA 1769
134 .....SerThrLysPheGluAspTyrAlaMetSerProArgIleThrG1 148
1768 AAGATACGCTGGAACCTGATACATATATCAATGATTAATAAACAGATGA 1719
```

```

148 yaspgluleuasnlglyleuarglavalmettyrasnlglyhispro 165
1718 cagcccaattgatacaca...tcagatgacagatgacagtaagaag 1672
165 lnsnlaasnlglyleuarglyleuarglyleuasnlailearg 181
1671 aatataaacaactgttaagaaactgtatgaatgaagaaactg 1622
182 valthcgluaalavaltrpytyrtyrserasphasnlaiprolesers 198
1621 caatcacaagatcaccattt.....tcaaaagtrgattgaaga 1581
198 nproasglsuserphelysarglsusergsuserasnleuvalserthrs 215
1580 nratcgtaagcaattaacagaatcactgcagatgacgttgcaaaataca 1531
215 ergluleuserleuemetarglnalaleuylsglnleuileasproasn 231
1530 aagacttacaaagatgcgctaaatagcattaaa...atgaacaatgcctcaa 1484
232 leuallathlysmetprolysglnvalproasphasnleuileas 248
1483 ttgcgtgaaacttaagaaacaaacttcattgatgat.....at 1446
248 epheglsuserglsaplyslysglysplys.....tyrasnlysglytyrg 263
1445 tgcacaaagaaacctgatacagatcacacatttatctatnaacatgctcaaac 1396
263 lnsnleuuserglslyleuvalprothrlys.....pro 275
1395 aagactttatagctcaggttaaatgacagatgaagctaaataatgaaga 1346
276 prothrproglyasproprometproproasnlglnproglinthrse 292
1345 gcaattgtcaaaagacaaacgtttataaaatgaatataatttgaaaaa 1296
292 valleuilearglystyralllelglyaspyrtyrserlyseuileug 309
1295 accgttaagcagaacacatttaattaacagattgacataacacagttcgcc 1246
309 lvalathrleuagleuthrlygaspasnvalasnserpheglnalialarg 325
1245 aagacacacaaagattgattgaatgaatgctgcaaaagtcacacgtaacgaa 1196
326 valphaseraserasnspilleglygluarglleuileuserasglyth 342
1195 acgatttaaaagatgatttaataca..... 1169
342 rtyrthrleuthrleuasnserproalaglytyrserllealaglup 359
1168 .....ttaaactgttcacacagatcct..... 1148
359 roilerhrphelysvalgluala.....glylysvaltyr 370
1147 .....catttttaattttgaaggcgacatttaaaatttaattgtaaaaaatat 1103
371 thrilleleasp...glylsglnllegluasproasnlysgluileva 386
1102 gacatttaagagataaagtttcaactgatacttaacaaagaa..... 1058
386 lgluprottyrservalglualatyrasnaspheglugluhpheeserval 403
1057 .....tataaagttgacgacat.....ggcgttg 1033
403 eutthrthrlnasntrtyralalyspherityralalysasnlysasngly 419
1032 ctaaaattgaaaaagatgctgagaaacattttcttaaaaaagataaacaatg 983
420 serseglvalvaltyrcyspheasnalaspleuylserserproproas 436
982 cattttcaaatgttatttt...ggacaaagcaaatgctgaacatgaaccnaa 936

```

```

436 psergluaspglyglylysthrmetthrproasphethrthrlygluval 453
935 t.....gataagaaagcaacagctgtgtgattgaacattgaacatca 893
453 allystyrthrhisllealaglyargaspheuleyslysthrvallys 469
892 .....aaccttgatggcgc...ttatcgaaagatgcatttaaac 857
470 proargaspthrproaspthrphelyslysnlelelysnvalil 486
856 cagcaatgtgagtcattatctagcttgaatgcccattatataaatgtaac 807
486 eglu...lysglytyrarggluyls..... 493
806 agattacaaaagcgacagaaagatmaaccatttcgacaaacagtttaattg 757
494 .....glylnalalleglutyrserglyleuthrthrln 506
756 atatgatggttgacacaaagtatc.....aatgcattgaa 722
507 leuargalalathrleuallaleuallaleuvaltyrtyrphethraspseralagl 523
721 agtttcaaaagcagataaagtagctgtgattacatcaaatgattcaatgca 672
523 uluasaplysasplysleu..... 529
671 agaaactcagacaaactgattgatgacatttttaatacaaaaagaaata 622
530 ....lysaspyrtnlsglypheglyasphetasnspserthrleuala 544
621 caacaaataatmaagaaagatattttccaaagctgattgatcgttaagaaac 572
545 valalalyslleuvalglutyrallaglnaspsersasproproglule 561
571 gtttaaaaagacttttgctgaagagccacaaagaaacaaatgatgaag 522
561 uthraspleuasphethrleuproasnasn..... 572
521 caaaaatgatgaatttaaatcagatgcttcacaaatttaagatgaagaattta 472
573 .....lystyglnserleuileglythrlntrphlspro 584
471 gtagaattttgaaagagacgacgaattgctatcagatatacagaagatca 422
585 gluaspleuvalaspilelleargmetgluaspylslysgluvalillepr 601
421 aaatcaattgcagattctggt.....agtgccaatttaaatca 384
601 ovalthrhisasnleuthrleuarglysthrvalthrlyleuala...g 617
383 agtcgacaaatgaatgtagat...aagctacatgcgacaggtgcagcatttag 337
617 lysapargthrlyaspphehispheglutllegluleuylsasasnlyls 633
336 gcgtaagcgttaacgatttaaatcgaatgcaaaatggctataaaacgataaagat 287
634 glnleuileuuserglnthrvallys..... 642
286 aatgagttgttcgctaaagaaatttaaaaaagatttaacaaattcttaaaagat 237
643 ....thasplysthr 646
236 gccgcacagacaaac 221

```

seq_name: /cgml_7/ptodata/2/pna/US08_NEW_COMB.seq:US-08-956-171C-264

```

seq_documentation_block:
; Sequence 264, Application US/08956171C
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
;           Gil H. Choi
;           Patrick S. Dillon
;           Craig A. Rosen

```


US-08-956-171C-168

alignment_scores:

Quality: 128.50 Length: 734
 Ratio: 0.377 Gaps: 37
 Percent Similarity: 46.458 Percent Identity: 18.601

alignment_block:

US-09-494-297-2 x US-08-956-171C-168/rev ..

Align seg 1/1 to reverse of: US-08-956-171C-168 from: 1 to: 7963

```

101 SerArgSerTyrGlnValTyrCysPheAsnLeuLysLysAlaPheProLe 117
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7385 ACMAATCAATCAATTTGTTAT...TTCAACAATAATTAATCTTTATTC.... 7344
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
117 uGlySerAspSerSerValLysLysTyrPtyLysLysHisAspGlyLys 134
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7343 .....CAGCATGCGCAACAATGCTTGAATTCACCTGACAGACAGCA 7304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
134 eRhrLysPheGlnAspTyrAlaMetSerProArgLleRhrGlyAspGlu 150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7303 ATGAAAAATGCGGTATTAATGCTGAAAAATACCAAACTGTTACAAATGAA 7254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
151 LeuAsnGlnLysLeuArgAlaValMetLysAsnGlnHisProGlnAsnAl 167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7253 TTTAAATCACT...TTAAATGCTCAATTA.....GAAGAAAATTA 7219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
167 aAsnGlyLleMetGluGlnLeuGluProLeuAsnAlaIleArgValTyrG 184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7218 TAAAGAATTTACAGACACTATTCGATTTTATCAATGTT..... 7182
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
184 lngAlaValAlTrrPtyTyrSerAspAsnAlaProIleSerAsnProAsp 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7181 .....GGTACTGAA 7173
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
201 GluSerPheLysArgGluSerGluSerAsnLeuValSerThrSerGlnLe 217
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7172 GAGATTTTTCACCAATCATGAGATTAACCAAACTTATCTACTAATAATT 7123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
217 uSerLeuMetArgGlnAlaLeuLysGlnLeuLleAspProAsnLeuAlaT 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7122 GAGTCGTTTAAATGATTTAACTAAATATCTTGAGACCAAAACATATTCTT 7073
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
234 hRlysMetProLysGlnValProAspAspPheGlnLeuSerIlePheGlu 250
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7072 ATGATTAAGTTCTAGTTTAAGTAAAAAACAACTGCACAACTT...GAA 7026
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
251 SerGluAsp.....LysGlyAspLysTyrAsnLysGlyTyr 262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7025 GAGGAAGATCATTTATTTGGCTACTCAAGTTGACGAATATTAATGAGCAATA 6976
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
262 rGlnAsnLeuLeuSerGlyGlyLeuValProThrLysProThrProG 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6975 TCTTGAATGCAACGA.....CAAGTCA 6953
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
279 LysAspProPheMetProAsnGlnProGlnThrSerValLeuIle 295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6952 GTGATTTAAAGTCACAAATCAATCAATGCAAACTGATACAAACCTTGCT 6903
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
296 Arg.....LysTyrAlaIleGlyAspTyrSe 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6902 AATTTAAGACATGAAATATCACAGTCTTAAAAATCAACTTAATGATATGCG 6853
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
304 rLysLeuLeuGlnGlyAlaThrLeu.....GlnLeuThrGlyA 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6852 AAAA.....GATGGCGCAAGTTTAAGTATTACAAAGTTTAGTGTGATG 6809
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
317 sPAsnValAsnSerPheGlnAlaArgValPheSerSerAsnAspIleGly 333
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6808 AACCATTAACAATAATTAAGATAACGTTTG...CTCAAGTTATTATAT 6762
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

334 GluArgIleGlu.....LeuSerAspGlyThrTyrThrLeuTh 346
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6761 GAGCGGTAGAAATATTGACACATTTAACAGATGCGACATATGATGAT 6712
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
346 rGlnLeuAsnSerProAlaGlyTyrSerIleAlaGluProIleThrPhe 363
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6711 TAACATTAATGAA.....GATCAATTAACGGTGA 6683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
363 ySValGlnAlaGlyLysValTyrThrIleIleAspGlyLysGlnIleGlu 379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6682 AACATGTTAATGGTCAATTAAT.....GATCCTGTTGAACCTAGT 6642
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
380 AsnProAsnLysGlnIleVal..... 386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6641 CAATCTACAAAGAAATTAATTGATGACCTTACGTATCACTTAATTA 6592
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
387 .....GluProTyrSerValGlnAlaTyrAsnAspPheG 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6591 GGTACTAGACCATATTTATCGCTGCCATTAATGTTGATGATGATG 6542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
398 lngLupheSerValLeuThrThrGlnAsnTyrAlaLysPheTyrTyrAla 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6541 TTCAATTTGATAAAAAACGTACTGAAAAAATGTAATTAATTACGATCA 6492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
415 LysAsnLysAsnGlySerSerGlnValAl.....Ty 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6491 TTATCAGAACCTA...TCAGTACTTACTTTCATGCTGTAAGATAATA 6443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
425 rCysPhe..... 427
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6442 TTGTTCCATCAAAAGAAATGATTAATTAACAAATAGAGGAAGCGCG 6393
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
428 .....AsnAlaAspLeuLysSerProProAspSerGluAsp... 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6392 AAACGATGAGAAATATAGAGATCTAAATCCCGAGATTCAGTATCNC 6343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
440 .....GlyGlyLysThrMetThrProAs 447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6342 TTTTCTTAGTCATAAAGCTACACAGGCGTAAACAGCAAGTAAGA 6293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
447 pPheThrThr.....GlyGluVal.....LysT 455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6292 TTATATGACATTTACATTTGCAAGATTAAGTGGTGAATTAAGAGGAAT 6243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
455 yTrhHisIleAlaGlyArgAspLeuPheLysTyrThrValLysProArg 471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6242 TTTGGACGCGTCAAAAAATGATATG.....GCAACATCAACCGTGA 6199
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
472 AspThrAspProAspThrPheLeuLysHisIleLysLysValIleGlyL 488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6198 GAA.....ATTGTCATGTT..... 6184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
488 sGlyTyrArgGluLysGlyGlnAlaIleGlnTyrSerGly.....L 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6183 .....AAAGTGCATCATTAATACTATCGCGGAATAAACA 6147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
502 eutRhrGluThrGlnLeuArgAlaAlaThr.....GlnLeuAlaIle 515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6146 TGAAGTCACAAATTAAGTACAGCACTGAAAGATTAATTAATAACA 6097
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
516 TyrTyrPheThrAspSerAlaGlnLeuAspLysAspLysLysAsp... 531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6096 GAACAATTTGTAGATGGTGCACCTTTATCACCGGACAGAAATACAAAGA 6047
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
532 .....TyrHisGlyPheGlyAspMetAsnAspSerThrLeu...AlaValA 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6046 GATTTCCATATTTTGGTACAGATATTGAAAAATGCTAATTTACAGCTATCA 5997
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
546 lAlysIleLeuValGluTyrAlaGlnAspSerAsnProProGlnLeuThr 562
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
5996 CACGTCATTTATGAAAAAATATCAAGAACGA..... 5965
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
563 AspLeuAspPhePhe.....IleProAsnAsnAsnLysTyrGlnSe 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

5964 ..... TTTTACACATATTCACAGCTGCTAGTCTGATCTCATCATTAAGT 5924
576 rleuilegilythrhrlntrphs.....ProgluA 586
      ||| .....|||
5923 TGCAGATGCGCTTAACGTCATCATGATTTACAGCATGTTCAGTATTCGAAGAT 5874
586 spdeuValasprllellearmetGuasapylslygluValile..... 600
      :.....|
5873 CAATTTGTGCATTTATATCCATGTTAAACAAAGCTTTGTATATAGTCGT 5824
601 .....ProValrh 603
5823 ATTATTTTCATGATATTGCTAAAGCTTTAGACATTCAGATGGTCCGTGTC 5774
603 rhtsasneuthrleuarglysthValthrgly.....leuAlag 617
      :||| :||| :||| :||| :|||
5773 GACGTCGTATACAGTCGAAGCACTTATTTAGCACACATCTCGATTGCGA 5724
617 lVAspgrthrlyAsprhenshpheguillegluIleuIlelyAsasAsnlys 633
      :||| :||| :||| :||| :|||
5723 GGTGATGAATAGTCTGAAGCAGCTCGTGAATTTAAACATTGAAGGAGAAAGA 5674
634 GlnGluIleuIleSerGlnthrVallysthAsplysthAsnleuGlnph 650
      ||||| :||| :||| :||| :|||
5673 ATCATGTTGTTAAGACATATGATTTTATCTCTCACTGTAAGTTTAGCTA 5624
650 elYAspGlylysalatrhIleasnleuIlelyshsGlyIleuSerleuThrL 667
      :||| :||| :||| :||| :|||
5623 TGGTCTCCAAAGACGCCATAC...TTAAAGAAACAGCAAAATTTTA... 5581
667 euGlnGlyleuProGluGlyTrSerTryleuVallysgluIthrAspser 683
5580 .....TGCATATAC..... 5572
684 GlnGlyTrLylyVallyAsasnSerGlnGluValAlaasnAlatrhVa 700
      :||| :||| :||| :||| :|||
5571 GATTAATATGATGCTAGACATGCAATATGTTTGAAGAAAGCATATTAAGAAAC 5522
700 lSerlysthGlyIlethrSerAspGluThrleuAlaIlehuGlnAsnL 717
5521 TGCACAAG...GGTCAGTTTACAGATAAATAATTTGGCTTGAATAATCTA 5475
717 ys 717
      :|||
5474 GA 5473

```

```

seq_name: /cgnl_7/ptodata/2/pna/PCT_NEW_COMB.seq:PCT-US01-03782A-2
seq_documentation_block:
  Sequence 2, Application PC/TUS0103782A
  GENERAL INFORMATION:
  APPLICANT: Hyseq, Inc.
  APPLICANT: Ford, John E et al
  TITLE OF INVENTION: Novel Bone Marrow Nucleic Acids and Polypep
  FILE REFERENCE: 21372-040
  CURRENT APPLICATION NUMBER: PCT/US01/03782A
  CURRENT FILING DATE: 2001-02-05
  PRIOR APPLICATION NUMBER: 09/496,914
  PRIOR FILING DATE: 2000-02-03
  PRIOR APPLICATION NUMBER: 09/598,075
  PRIOR FILING DATE: 2000-06-20
  PRIOR APPLICATION NUMBER: 09/620,325
  PRIOR FILING DATE: 2000-07-19
  PRIOR APPLICATION NUMBER: 60/250,583
  PRIOR FILING DATE: 2000-11-30
  NUMBER OF SEQ ID NOS: 386
  SOFTWARE: Custom
  SEQ ID NO 2
  LENGTH: 16122
  TYPE: DNA
  ORGANISM: Homo sapiens
  PCT-US01-03782A-2

```

```
alignment_scores:      Length: 7222
      Quality:          Gaps: 36
      Ratio: 0.372      Percent Identity: 20.637
Percent Similarity: 47.230
```

```
alignment_block:
US-09-494-297-2 x PCT-US01-03782A-2 ..
```

Align seg 1/1 to: PCT-US01-03782A-2 from: 1 to: 16122

[illegible]

```

7609 AAAGACTCCAGACATCAATCAG..... 7632
      |||:|||||
306 uleugluGlyAlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheG 323
      :|||:|||||
7633 .....AATGGCTTGAGCTGACT...GACAACTCACTCACTCCGTT 7669
      :|||:|||||
323 LnaIaArgValPheSerSerAsnAspIleGlyGluArgIleGluLeuSer 339
      :|||:|||||
7670 CCAGCCCAATTT.....GACCAAGCTATTGTAAAGACACCCAGTACAG 7713
      :|||:|||||
340 AspIleThrTyThrLeuThrGlnLeuAsnSerProAlaGly.....Ty 354
      :|||:|||||
7714 GAATGCTCCAGACTATATCAGAAAGGTGAGGCGACTGGACAAACGGCT 7763
      :|||:|||||
354 rSerIleAlaGlnProIleThrPheLysValGluAlaGlyLysValTyT 371
      :|||:|||||
7764 GAGTGCACAGCTATACAGCACCAACAGAGGCTGTAAAG..... 7806
      :|||:|||||
371 hrIleIleAspGlyLysGlnIleGluAsnProAsnLys..... 383
      :|||:|||||
7807 .....CAGCAATTTGAGAGACACACTGAATTCGATCTGAC 7842
      :|||:|||||
384 ...GluIleValGluProTySerValGluAlaTyAsnAspPheGluI 399
      :|||:|||||
7843 TTGGAGCAGTTAGACACAGAGTTAAGAGGCTCAGACACTGTCGATGA 7892
      :|||:|||||
399 uPheSerValLeuThrThrGlnAsnTyAlaLysPheTyThrAlaLysA 416
      :|||:|||||
7893 ACTCTCAGTGTCTATGTGTGAGCAGTACCTCAAGATGAAGTGAAGAAG 7942
      :|||:|||||
416 sn.....LysAsnGlySerSer 421
      :|||:|||||
7943 GTTTGGAGACAGTTGCCCTGCTCCAGAGTTTGAAGAGACCTTGACGC 7992
      :|||:|||||
422 GlnValValTyCysPheAsnAlaAspLeuLysSerProProAspSerG 438
      :|||:|||||
7993 GATCGCATTAACAGACTCCAGGACCTTGTCCAGCACCCAGCGATTCCA 8042
      :|||:|||||
438 u.....AspGlyLysTyThrMetThrProAspPheThrThrGlyG 452
      :|||:|||||
8043 GCAAATGTTTATGATGATGAGGACCTGTTGATGATTAACAAGCCAGC 8092
      :|||:|||||
452 LuValLysTyThrHisIleAlaGlyArgAspLeuPheLysTyThrVal 468
      :|||:|||||
8093 AACCAAAAACCTGCCCAATTTCTCAAAA..... 8121
      :|||:|||||
469 LysProAlaArgAspThrAspProAspThrPheLeuLysHisIleLys 485
      :|||:|||||
8122 .....TTGGAGCGGCTACAGTCCA 8141
      :|||:|||||
485 LileGluLysGlyTyArgGluLysGlyGlnAlaIleGluTySerGlyL 502
      :|||:|||||
8142 GCTACAGAGAGATGAAGGTTTCAGAAAAGCTTAATCAACACAGTGGCT 8191
      :|||:|||||
502 euThrGluThrGlnLeuArgAlaAlaThrGlnLeuAlaIle-TyTyTPh 518
      :|||:|||||
8192 CCTATGAGGTCAATGTGCTCAAGGGAATCTCACTTCTTGTACTACT 8241
      :|||:|||||
518 eThrAspSerIleGluLeuAspLysAspLysLeuLysAspTyThrHisG 535
      :|||:|||||
8242 CCTG.....GAGAAAGAGAAAGAGACTTACAAA 8270
      :|||:|||||
535 heGlyAspMetAsnAspSerThrLeuAlaValAlaLysIleLeuValGlu 551
      :|||:|||||
8271 CCAGTTGGTTGAGCTCAAAAAC.....ATTGGGAAGAGCTTAGTAA 8314
      :|||:|||||
552 TyAlaGlnAspSerAsnProProGlnLeuThrAspLeuAspPhePheI 568
      :|||:|||||
8315 AAAGTCGAGACAGCAATCCAG.....CTCAAGATTTGTAT 8351
      :|||:|||||
568 eProAsnAsnAsnLysTyGlnSerLeuIleGlyThrGlnThrPheProG 585
      :|||:|||||

```

```

8352 GCAGAAAGCTCAGAAATAT.....CAGTGCAGTGG 8383
      :|||:|||||
585 LuAspLeuValAspIleIle.....ArgMetGluAspLys 596
      :|||:|||||
8384 AAGACCTTGTCATGATGATGATGATGATGATGATGATGATGATGATG 8433
      :|||:|||||
597 LysGluValIle...ProValThrHisAsnLeuThrLeuArgLysTyTh 612
      :|||:|||||
8434 CGAGTCACTGTGATCCAGTCCAGCTAGAGTCCAGTCTCTTAAGATCAA 8483
      :|||:|||||
612 LThrGlyLeuAlaGlyAspArgThrLysAspPheHisPheGluIleGlu 629
      :|||:|||||
8484 GCGTATGCTG.....AATGAGTGAAGAGAGCGCGCTCCCTGCGTGA 8527
      :|||:|||||
629 euLysAsnAsnLysGlnGluLeuLeuSerGlnThrValLysThrAspLys 645
      :|||:|||||
8528 TATGAAATAGTGTGCTGACATCTGATCAATCTGTGAAGAGAGATGAG 8577
      :|||:|||||
646 ThrAsnLeuGluPheLysAspGlyLysAlaThrIleAsn..... 658
      :|||:|||||
8578 GATCGAATC.....CGGATGAGAGAGGCTGGGATCAACAGAAATGGA 8621
      :|||:|||||
659 .....LeuLysHisGly.....GluSerLeuT 666
      :|||:|||||
8622 TGCCTTACAGAGAGCTGCGAGGCCAAACAGGCTCAGCTCAGAAATGA 8671
      :|||:|||||
666 hr.....LeuGlnGlyLeuProGluGlyTyTySerTyThrValLysGlu 680
      :|||:|||||
8672 CTCAGAGCTCAGGAGAGTCCAGAGAAAGCTTAAATATTAAGAAAGAG 8721
      :|||:|||||
681 ThrAspSerGluGlyTyThrLysValLysVal.....AsnSerG 693
      :|||:|||||
8722 GTTGAAGAGGCCAAACCAACTTGAGATCTTGAATGCTGTGGTCTTCA 8771
      :|||:|||||
693 nGluValAlaAsnAlaThrValSerLysThrGlyIleThrSerAspGlu 710
      :|||:|||||
8772 AGCCTGTAACAAAGAACTTGAGAAAGCTTAAGAGCTCAACAGGAAGTC 8821
      :|||:|||||
710 hrLeuAlaPheGlu 714
      :|||:|||||
8822 TGCAGGCCCTGAGAG 8835
      :|||:|||||

```

seq_name: /cgnl_7/prodate/2/pna/PCT_NEW_COMB.seq:PCT-US01-03782A-190

seq_documentation_block:

```

; Sequence 190, Application PC/TUS0103782A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Bone Marrow Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-040
; CURRENT APPLICATION NUMBER: PCT/US01/03782A
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/250,583
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 386
; SOFTWARE: Custom
; SEQ ID NO 190
; LENGTH: 16341
; TYPE: DNA
; ORGANISM: Homo sapiens
; PCT-US01-03782A-190

```

alignment_scores:

```

Quality: 127.00      Length: 722
Ratio: 0.372         Gaps: 36

```

Percent Similarity: 47.230 Percent Identity: 20.637

Alignment block:

US-09-494-297-2 x PCT-US01-03782A-190 ..

Align seg 1/1 to: PCT-US01-03782A-190 from: 1 to: 16341

```

84 TyrTyrLysGlnPheArgValAlaHisAspLeuArgValAsnLeuGluG1 100
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
786 TATAGAAACTAGGAGGAGTACTTCATGACGCAAGCAAGCAAGCTTCAGGC 7235
   : : : : : : : : : : : : : : : : : : : : : : : : : :
100 Y.....SerArgSerTyrGlnValTyrCysPheAsnLeuLysLysAla 115
   : : : : : : : : : : : : : : : : : : : : : : : : : :
736 TATCCTCAACAGAAATGAGAGGAGTTTCAC.....AAGAGGCA. 7272
   : : : : : : : : : : : : : : : : : : : : : : : : : :
115 heProLeuGlySerAspSerSerValLysLysTyrTyrLysHisAsp 131
   : : : : : : : : : : : : : : : : : : : : : : : : : :
7273 .....AACTGTGTGTGTCAGTGGCTGGAATCAAAAGAG 7305
   : : : : : : : : : : : : : : : : : : : : : : : : : :
132 GlyLleSerThrLysPheGluAspTyrAlaMetSerProAlaIleThrG1 148
   : : : : : : : : : : : : : : : : : : : : : : : : : :
7306 GAATC...CTGAATCCATGTGATGCCATCTCATCTCCAAACCAAGACA 7352
   : : : : : : : : : : : : : : : : : : : : : : : : : :
148 YAsp.....GluLeuAsnGlnLysLeuArgAlaValMet 160
   : : : : : : : : : : : : : : : : : : : : : : : : : :
7353 AACAGTGAAGCCCAAGCTGAATCTAACAAAGCCTCTCGGTGAGTGTG. 7401
   : : : : : : : : : : : : : : : : : : : : : : : : : :
160 YrAsnGlyHisProGlnAsnAlaAsnGlyIleMetGluGlyLeuGluPro 176
   : : : : : : : : : : : : : : : : : : : : : : : : : :
7402 .....GAACAGAAATTCCTCAAAATTCAAAGTAAGAGAACCC 7440
   : : : : : : : : : : : : : : : : : : : : : : : : : :
177 LeuAsnAlaIleArgValThrGlnGluAlaValTrrTyrTyrSerAspAs 193
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7441 CTGGCTGATTAAGTACTGTCACATAT..... 7464
   : : : : : : : : : : : : : : : : : : : : : : : : : :
193 nAlaProIleSerAsnProAspGluSerPheLysArg..... 205
   : : : : : : : : : : : : : : : : : : : : : : : : : :
7465 ....CCCAACTCACAGCAAGCAAGAAATTTGGAAGAAATTCAGAGAACA 7510
   : : : : : : : : : : : : : : : : : : : : : : : : : :
205 ..... 205
7511 TCAATTCCTCCGATGGAAAGGCCACTGAGTTACTGTGGCTGGCAAGG 7560
   : : : : : : : : : : : : : : : : : : : : : : : : : :
206 .....GluSerGluSerAsnLeu.....ValSerThrSe 215
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
7561 CAGCTAGAGGAAATCTGCAATCATCTGGCCCTGCTCCAGGCTGCAGATC 7610
   : : : : : : : : : : : : : : : : : : : : : : : : : :
215 rGlnLeu.....SerLeuMetArgGlnAlaLeuL 225
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
7611 CCAGCTCCAGCCGCTGGCTGATGAGAAAGAACTATGATGGAGTGGCTGG 7660
   : : : : : : : : : : : : : : : : : : : : : : : : : :
225 ysnGlnLeu...IleAspProAsnLeuAlaThrLysMetProLysGlnVal 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
7661 GGCCCTGTCTATTGACCCCAACATGTTGAATGACAAAGCAACAGGCTC 7710
   : : : : : : : : : : : : : : : : : : : : : : : : : :
241 ProAspAspPheGlnLeuSerIlePheGluSer...GluAspLysGlyAs 256
   : : : : : : : : : : : : : : : : : : : : : : : : : :
7711 .....CAGTTTATGCTTAAAGCAATTTCAACACAGCAACAGCATGA 7754
   : : : : : : : : : : : : : : : : : : : : : : : : : :
256 pLysTyrAsnLysGlyTyrGlnAsnLeuLeuSerGlyLysLeuValProT 273
   : : : : : : : : : : : : : : : : : : : : : : : : : :
7755 GCAACTGAATGAGCAGCTCAGGCAATCTTAACAGC..... 7791
   : : : : : : : : : : : : : : : : : : : : : : : : : :
273 hrLysProProThrProGlyAspProPheMetProProAsnGlnProGln 289
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
7792 .....CTGAGATGTCTCTCTGTCACACCAAGTACAG 7837
   : : : : : : : : : : : : : : : : : : : : : : : : : :
290 ThrThrSerValLeuIleArgLysTyrAlaIleGlyAspTyrSerLysLe 306
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
7828 AAAGAACTCCAGAGCATCAATCAG..... 7851
   : : : : : : : : : : : : : : : : : : : : : : : : : :
306 uLeuGluGlyAlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheG 323
   : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

7852 .....AAATGGTTGAGCTGACT...GACAACTCAACTCCGCTT 7888
323 InAlaArgValPheSerSerAsnAspIleGlyGluArgIleGluLeuSer 339
   : : : : : : : : : : : : : : : : : : : : : : : : : :
7889 CCAGCCAAATTT.....GACCACCTATTGTTAAGACACCACTACAGC 7932
   : : : : : : : : : : : : : : : : : : : : : : : : : :
340 AspGlyThrTyrThrLeuThrGluLeuAsnSerProAlaGly.....Ty 354
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
7933 GAATGCTCCAGCACTTATATGACAGAGGTGAGGCGACGTGGACAACGGCT 7982
   : : : : : : : : : : : : : : : : : : : : : : : : : :
354 rSerIleAlaGluProIleThrPheLysValGluAlaGlyLysValTyrT 371
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
7983 GAGTGTCCAGTCAGCTATCAACCAACCAAGAGGCTGTAAAG..... 8025
   : : : : : : : : : : : : : : : : : : : : : : : : : :
371 hrIleIleAspGlyLysGlnIleGluAsnProAsnLys..... 383
   : : : : : : : : : : : : : : : : : : : : : : : : : :
8026 .....CAGCAATGTGAAGACACCACTGAATTCGATCTGCAC 8061
   : : : : : : : : : : : : : : : : : : : : : : : : : :
384 ..GluIleValGluProTyrSerValGluAlaTyrAsnAspPheGluG1 399
   ||| : : : : : : : : : : : : : : : : : : : : : : : :
8062 TTGGAGCAGTTAGACCAAGAGCTTAAGAGGCTCAGACACTGTGCATGA 8111
   : : : : : : : : : : : : : : : : : : : : : : : : : :
399 uPheSerValLeuThrThrGlnAsnTyrAlaLysPheTyrTyrAlaLysA 416
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
8112 ACTCTGAGTCTCATTTGTGTGAGCAGTACTCAAGATGAACGTGAAGAGC 8161
   : : : : : : : : : : : : : : : : : : : : : : : : : :
416 sn.....LysAsnGlySerSer 421
   : : : : : : : : : : : : : : : : : : : : : : : : : :
8162 GTTTGGAGACAGTGGCCCTGCTCTCAAGGTTTGAAGACCTTGCAAGC 8211
   : : : : : : : : : : : : : : : : : : : : : : : : : :
422 GlnValValTyrCysPheAsnAlaAspLeuLysSerProAspSerG1 438
   : : : : : : : : : : : : : : : : : : : : : : : : : :
8212 GATCGCACTTAACAGACATCCAGCAGCTCTTGCCAGCACCCAGAGTTCA 8261
   : : : : : : : : : : : : : : : : : : : : : : : : : :
438 u.....AspGlyGlyThrMetThrProAspPheThrGlyG 452
   : : : : : : : : : : : : : : : : : : : : : : : : : :
8262 GCAATGTTTGTGATGATGAGACCTGTTGATGATTAACAAAGCAGC 8311
   : : : : : : : : : : : : : : : : : : : : : : : : : :
452 LuValLysTyrThrHisIleAlaGlyArgAspLeuPheLysTyrThrVal 468
   : : : : : : : : : : : : : : : : : : : : : : : : : :
8312 AAGCAAAAACCTGCCCAATTTCTGCAAA..... 8340
   : : : : : : : : : : : : : : : : : : : : : : : : : :
469 LysProArgAspThrAspProAspThrPheLeuLysHisIleLysVal 485
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
8341 .....TTGAGCGGCTACAGTCTCA 8360
   : : : : : : : : : : : : : : : : : : : : : : : : : :
485 IleGluLysGlyTyrArgGluLysGlyGlnAlaIleGluTyrSerGlyL 502
   : : : : : : : : : : : : : : : : : : : : : : : : : :
8361 GCTACAGAGAGATGAGAGTTTCAGAAAGCTTAATCAACACAGTGGCT 8410
   : : : : : : : : : : : : : : : : : : : : : : : : : :
502 euThrGluThrGlnLeuArgAlaAlaThrGlnLeuAlaIleTyrTyrPh 518
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
8411 CCTATGAGGTATTTGTGCTCAAGGGAATCTCTACTCTTCTGTAAGCT 8460
   : : : : : : : : : : : : : : : : : : : : : : : : : :
518 eThrAspSerAlaGluLeuAspLysAspLysLeuLysAspTyrHisGlyP 535
   : : : : : : : : : : : : : : : : : : : : : : : : : :
8461 CCTG.....GAGAAGAGAAAGACACTACAAAA 8489
   : : : : : : : : : : : : : : : : : : : : : : : : : :
535 heGlyAspMetAsnAspSerThrLeuAlaValAlaLysIleLeuValGlu 551
   : : : : : : : : : : : : : : : : : : : : : : : : : :
8490 CCAATGTTGAGCTCAAAAAAC.....ATTGGAAAGAGCTTGTGTAAA 8533
   : : : : : : : : : : : : : : : : : : : : : : : : : :
552 TyrAlaGlnAspSerAsnProProGlnLeuThrAspLeuAspPhePheI 568
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
8534 AAACCTGCAAGACAGCAATCCAGC.....CTCAAGATTGTAT 8570
   : : : : : : : : : : : : : : : : : : : : : : : : : :
568 eProAsnAsnAsnLysTyrGlnSerLeuIleGlyThrGlnThrPheProG 585
   : : : : : : : : : : : : : : : : : : : : : : : : : :
8571 GCAGAAAGCTCAGAAATAT.....CAGTGGCATGTGG 8602
   : : : : : : : : : : : : : : : : : : : : : : : : : :
585 LuAspLeuValAspIleLe.....ArgMetGluAspLys 596
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
8603 AAGACCTTGTGCTCATGATGATGAAGATTGTAAGACTAGATGTGAGATTG 8652
   : : : : : : : : : : : : : : : : : : : : : : : : : :

```


